

RECEIVED

SEQUENCE LISTING

APR 1 2 2002

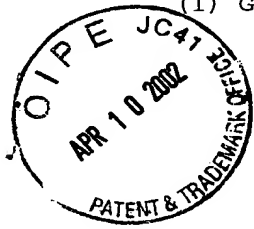
TECH CENTER 1600/2900

(1) GENERAL INFORMATION:

- (i) APPLICANT: Jaye, Michael C.
Lynch, Kevin J.
Amin, Dilip V.
Doan, Kim-Anh T.
Marchadier, Dawn
Maugeais, Cyrille
Rader, Daniel J.
Krawiec, John A.
South, Victoria J.
- (ii) TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR EFFECTING THE LEVELS OF HIGH DENSITY LIPOPROTEIN (HDL) CHOLESTEROL AND APOLIPOPROTEIN AI, VERY LOW DENSITY LIPOPROTEIN (VLDL) CHOLESTEROL AND LOW DENSITY LIPOPROTEIN (LDL) CHOLESTEROL
- (iii) NUMBER OF SEQUENCES: 31
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Synnestvedt & Lechner LLP
(B) STREET: Suite 2600 Aramark Tower, 1101 Market Street
(C) CITY: Philadelphia
(D) STATE: PA
(E) COUNTRY: USA
(F) ZIP: 19107
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-Windows 95
(D) SOFTWARE: Corel WordPerfect 8.0 converted to ASCII
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 09/277,401
(B) FILING DATE: March 26, 1999
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Kelly, Ph.D., Patrick J.
(B) REGISTRATION NUMBER: 34,638
(C) REFERENCE/DOCKET NUMBER: 22,944-C USA
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (215) 923-4466
(B) TELEFAX: (215) 923-2189

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 367 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 22..180
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
- | | | | | | | | | | | | | | | | | |
|------------|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|
| GAATTCGGCT | TGATCAATCG | C | TTC | AAA | AAG | GGG | ATC | TGT | CTG | AGC | TGC | CGC | .51 | | | |
| | | | Phe | Lys | Lys | Gly | Ile | Cys | Leu | Ser | Cys | Arg | | | | |
| | | | 1 | | | | 5 | | | | | 10 | | | | |
| AAG | AAC | CGT | TGT | AAT | AGC | ATT | GGC | TAC | AAT | GCC | AAG | AAA | ATG | AGG | AAC | 99 |
| Lys | Asn | Arg | Cys | Asn | Ser | Ile | Gly | Tyr | Asn | Ala | Lys | Lys | Met | Arg | Asn | |
| | | | 15 | | | | 20 | | | | | | 25 | | | |



7'

[illegible]

(i) SEQUENCE CHARACTERISTICS:

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 312..1370

(xi) SEQUENCE DESCRIPTION: SEO ID NO:3:

GAATTCGGCT	TCTACTACTA	CTAGGCCACG	CGTCGCCTAG	TACGGGGGGG	GGGGGGGGGG	60
TCAGCGAGTC	CTTGCCTCCC	GGCGGCTCAG	GACGAGGGCA	GATCTCGTTC	TGGGGCAAGC	120
CGTTGACACT	CGCTCCCTGC	CACCGCCCGG	GCTCCGTGCC	GCCAAGTTTT	CATTTTCCAC	180
CTTCTCTGCC	TCCAGTCCCC	CAGCCCCTGG	CCGAGAGAAG	GGTCTTACCG	GCCGGGATTG	240
CTGGAAACAC	CAAGAGGTGG	TTTTTGTTTT	TTAAAACTTC	TGTTTCTTGG	GAGGGGGTGT	300
GGCGGGGCAG	G ATG AGC AAC TCC GTT	CCT CTG CTC TGT TTC TGG AGC CTC				350
	Met Ser Asn Ser Val	Pro Leu Leu Cys Phe Trp Ser Leu				
	1	5	10			
TGC TAT TGC TTT GCT GCG GGG AGC CCC GTA CCT TTT GGT CCA GAG GGA						398
Cys Tyr Cys Phe Ala Ala Gly Ser Pro Val Pro Phe Gly Pro Glu Gly						
	15	20	25			
CGG CTG GAA GAT AAG CTC CAC AAA CCC AAA GCT ACA CAG ACT GAG GTC						446
Arg Leu Glu Asp Lys Leu His Lys Pro Lys Ala Thr Gln Thr Glu Val						
	30	35	40			45
AAA CCA TCT GTG AGG TTT AAC CTC CGC ACC TCC AAG GAC CCA GAG CAT						494
Lys Pro Ser Val Arg Phe Asn Leu Arg Thr Ser Lys Asp Pro Glu His						
	50	55	60			
GAA GGA TGC TAC CTC TCC GTC GGC CAC AGC CAG CCC TTA GAA GAC TGC						542
Glu Gly Cys Tyr Leu Ser Val Gly His Ser Gln Pro Leu Glu Asp Cys						
	65	70	75			
AGT TTC AAC ATG ACA GCT AAA ACC TTT TTC ATC ATT CAC GGA TGG ACG						590
Ser Phe Asn Met Thr Ala Lys Thr Phe Phe Ile Ile His Gly Trp Thr						
	80	85	90			

ATG AGC GGT ATC TTT GAA AAC TGG CTG CAC AAA CTC GTG TCA GCC CTG 638
 Met Ser Gly Ile Phe Glu Asn Trp Leu His Lys Leu Val Ser Ala Leu
 95 100 105
 CAC ACA AGA GAG AAA GAC GCC AAT GTA GTT GTG GTT GAC TGG CTC CCC 686
 His Thr Arg Glu Lys Asp Ala Asn Val Val Val Val Asp Trp Leu Pro
 110 115 120 125
 CTG GCC CAC CAG CTT TAC ACG GAT GCG GTC AAT AAT ACC AGG GTG GTG 734
 Leu Ala His Gln Leu Tyr Thr Asp Ala Val Asn Asn Thr Arg Val Val
 130 135 140
 GGA CAC AGC ATT GCC AGG ATG CTC GAC TGG CTG CAG GAG AAG GAC GAT 782
 Gly His Ser Ile Ala Arg Met Leu Asp Trp Leu Gln Glu Lys Asp Asp
 145 150 155
 TTT TCT CTC GGG AAT GTC CAC TTG ATC GGC TAC AGC CTC GGA GCG CAC 830
 Phe Ser Leu Gly Asn Val His Leu Ile Gly Tyr Ser Leu Gly Ala His
 160 165 170
 GTG GCC GGG TAT GCA GGC AAC TTC GTG AAA GGA ACG GTG GGC CGA ATC 878
 Val Ala Gly Tyr Ala Gly Asn Phe Val Lys Gly Thr Val Gly Arg Ile
 175 180 185
 ACA GGT TTG GAT CCT GCC GGG CCC ATG TTT GAA GGG GCC GAC ATC CAC 926
 Thr Gly Leu Asp Pro Ala Gly Pro Met Phe Glu Gly Ala Asp Ile His
 190 195 200 205
 AAG AGG CTC TCT CCG GAC GAT GCA GAT TTT GTG GAT GTC CTC CAC ACC 974
 Lys Arg Leu Ser Pro Asp Asp Ala Asp Phe Val Asp Val Leu His Thr
 210 215 220
 TAC ACG CGT TCC TTC GGC TTG AGC ATT GGT ATT CAG ATG CCT GTG GGC 1022
 Tyr Thr Arg Ser Phe Gly Leu Ser Ile Gly Ile Gln Met Pro Val Gly
 225 230 235
 CAC ATT GAC ATC TAC CCC AAT GGG GGT GAC TTC CAG CCA GGC TGT GGA 1070
 His Ile Asp Ile Tyr Pro Asn Gly Gly Asp Phe Gln Pro Gly Cys Gly
 240 245 250
 CTC AAC GAT GTC TTG GGA TCA ATT GCA TAT GGA ACA ATC ACA GAG GTG 1118
 Leu Asn Asp Val Leu Gly Ser Ile Ala Tyr Gly Thr Ile Thr Glu Val
 255 260 265
 GTA AAA TGT GAG CAT GAG CGA GCC GTC CAC CTC TTT GTT GAC TCT CTG 1166
 Val Lys Cys Glu His Glu Arg Ala Val His Leu Phe Val Asp Ser Leu
 270 275 280 285
 GTG AAT CAG GAC AAG CCG AGT TTT GCC TTC CAG TGC ACT GAC TCC AAT 1214
 Val Asn Gln Asp Lys Pro Ser Phe Ala Phe Gln Cys Thr Asp Ser Asn
 290 295 300
 CGC TTC AAA AAG GGG ATC TGT CTG AGC TGC CGC AAG AAC CGT TGT AAT 1262
 Arg Phe Lys Lys Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn
 305 310 315
 AGC ATT GGC TAC AAT GCC AAG AAA ATG AGG AAC AAG AGG AAC AGC AAA 1310
 Ser Ile Gly Tyr Asn Ala Lys Lys Met Arg Asn Lys Arg Asn Ser Lys
 320 325 330
 ATG TAC CTA AAA ACC CGG GCA GGC ATG CCT TTC AGA GGT AAC CTT CAG 1358
 Met Tyr Leu Lys Thr Arg Ala Gly Met Pro Phe Arg Gly Asn Leu Gln
 335 340 345
 TCC CTG GAG TGT CAAGCCGAAT TC 1382
 Ser Leu Glu Cys
 350

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 353 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ser Asn Ser Val Pro Leu Leu Cys Phe Trp Ser Leu Cys Tyr Cys
 1 5 10 15

Phe Ala Ala Gly Ser Pro Val Pro Phe Gly Pro Glu Gly Arg Leu Glu
 20 25 30
 Asp Lys Leu His Lys Pro Lys Ala Thr Gln Thr Glu Val Lys Pro Ser
 35 40 45
 Val Arg Phe Asn Leu Arg Thr Ser Lys Asp Pro Glu His Glu Gly Cys
 50 55 60
 Tyr Leu Ser Val Gly His Ser Gln Pro Leu Glu Asp Cys Ser Phe Asn
 65 70 75 80
 Met Thr Ala Lys Thr Phe Phe Ile Ile His Gly Trp Thr Met Ser Gly
 85 90 95
 Ile Phe Glu Asn Trp Leu His Lys Leu Val Ser Ala Leu His Thr Arg
 100 105 110
 Glu Lys Asp Ala Asn Val Val Val Val Asp Trp Leu Pro Leu Ala His
 115 120 125
 Gln Leu Tyr Thr Asp Ala Val Asn Asn Thr Arg Val Val Gly His Ser
 130 135 140
 Ile Ala Arg Met Leu Asp Trp Leu Gln Glu Lys Asp Asp Phe Ser Leu
 145 150 155 160
 Gly Asn Val His Leu Ile Gly Tyr Ser Leu Gly Ala His Val Ala Gly
 165 170 175
 Tyr Ala Gly Asn Phe Val Lys Gly Thr Val Gly Arg Ile Thr Gly Leu
 180 185 190
 Asp Pro Ala Gly Pro Met Phe Glu Gly Ala Asp Ile His Lys Arg Leu
 195 200 205
 Ser Pro Asp Asp Ala Asp Phe Val Asp Val Leu His Thr Tyr Thr Arg
 210 215 220
 Ser Phe Gly Leu Ser Ile Gly Ile Gln Met Pro Val Gly His Ile Asp
 225 230 235 240
 Ile Tyr Pro Asn Gly Gly Asp Phe Gln Pro Gly Cys Gly Leu Asn Asp
 245 250 255
 Val Leu Gly Ser Ile Ala Tyr Gly Thr Ile Thr Glu Val Val Lys Cys
 260 265 270
 Glu His Glu Arg Ala Val His Leu Phe Val Asp Ser Leu Val Asn Gln
 275 280 285
 Asp Lys Pro Ser Phe Ala Phe Gln Cys Thr Asp Ser Asn Arg Phe Lys
 290 295 300
 Lys Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn Ser Ile Gly
 305 310 315 320
 Tyr Asn Ala Lys Lys Met Arg Asn Lys Arg Asn Ser Lys Met Tyr Leu
 325 330 335
 Lys Thr Arg Ala Gly Met Pro Phe Arg Gly Asn Leu Gln Ser Leu Glu
 340 345 350
 Cys

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1065 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1065

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATG AGC AAC TCC GTT CCT CTG CTC TGT TTC TGG AGC CTC TGC TAT TGC
 Met Ser Asn Ser Val Pro Leu Leu Cys Phe Trp Ser Leu Cys Tyr Cys
 1 5 10 15

48

TTT GCT GCG GGG AGC CCC GTA CCT TTT GGT CCA GAG GGA CGG CTG GAA

96

Phe	Ala	Ala	Gly	Ser	Pro	Val	Pro	Phe	Gly	Pro	Glu	Gly	Arg	Leu	Glu		
			20					25					30				
GAT	AAG	CTC	CAC	AAA	CCC	AAA	GCT	ACA	CAG	ACT	GAG	GTC	AAA	CCA	TCT	144	
Asp	Lys	Leu	His	Lys	Pro	Lys	Ala	Thr	Gln	Thr	Glu	Val	Lys	Pro	Ser		
		35					40					45					
GTG	AGG	TTT	AAC	CTC	CGC	ACC	TCC	AAG	GAC	CCA	GAG	CAT	GAA	GGA	TGC	192	
Val	Arg	Phe	Asn	Leu	Arg	Thr	Ser	Lys	Asp	Pro	Glu	His	Glu	Gly	Cys		
		50				55					60						
TAC	CTC	TCC	GTC	GGC	CAC	AGC	CAG	CCC	TTA	GAA	GAC	TGC	AGT	TTC	AAC	240	
Tyr	Leu	Ser	Val	Gly	His	Ser	Gln	Pro	Leu	Glu	Asp	Cys	Ser	Phe	Asn		
		65			70					75					80		
ATG	ACA	GCT	AAA	ACC	TTT	TTC	ATC	ATT	CAC	GGA	TGG	ACG	ATG	AGC	GGT	288	
Met	Thr	Ala	Lys	Thr	Phe	Phe	Ile	Ile	His	Gly	Trp	Thr	Met	Ser	Gly		
			85						90					95			
ATC	TTT	GAA	AAC	TGG	CTG	CAC	AAA	CTC	GTG	TCA	GCC	CTG	CAC	ACA	AGA	336	
Ile	Phe	Glu	Asn	Trp	Leu	His	Lys	Leu	Val	Ser	Ala	Leu	His	Thr	Arg		
			100					105						110			
GAG	AAA	GAC	GCC	AAT	GTA	GTT	GTG	GTT	GAC	TGG	CTC	CCC	CTG	GCC	CAC	384	
Glu	Lys	Asp	Ala	Asn	Val	Val	Val	Val	Asp	Trp	Leu	Pro	Leu	Ala	His		
		115					120						125				
CAG	CTT	TAC	ACG	GAT	GCG	GTC	AAT	AAT	ACC	AGG	GTG	GTG	GGA	CAC	AGC	432	
Gln	Leu	Tyr	Thr	Asp	Ala	Val	Asn	Asn	Thr	Arg	Val	Val	Gly	His	Ser		
		130				135							140				
ATT	GCC	AGG	ATG	CTC	GAC	TGG	CTG	CAG	GAG	AAG	GAC	GAT	TTT	TCT	CTC	480	
Ile	Ala	Arg	Met	Leu	Asp	Trp	Leu	Gln	Glu	Lys	Asp	Asp	Phe	Ser	Leu		
		145			150					155					160		
GGG	AAT	GTC	CAC	TTG	ATC	GGC	TAC	AGC	CTC	GGA	GCG	CAC	GTG	GCC	GGG	528	
Gly	Asn	Val	His	Leu	Ile	Gly	Tyr	Ser	Leu	Gly	Ala	His	Val	Ala	Gly		
				165					170					175			
TAT	GCA	GGC	AAC	TTC	GTG	AAA	GGA	ACG	GTG	GGC	CGA	ATC	ACA	GGT	TTG	576	
Tyr	Ala	Gly	Asn	Phe	Val	Lys	Gly	Thr	Val	Gly	Arg	Ile	Thr	Gly	Leu		
			180					185						190			
GAT	CCT	GCC	GGG	CCC	ATG	TTT	GAA	GGG	GCC	GAC	ATC	CAC	AAG	AGG	CTC	624	
Asp	Pro	Ala	Gly	Pro	Met	Phe	Glu	Gly	Ala	Asp	Ile	His	Lys	Arg	Leu		
		195					200						205				
TCT	CCG	GAC	GAT	GCA	GAT	TTT	GTG	GAT	GTC	CTC	CAC	ACC	TAC	ACG	CGT	672	
Ser	Pro	Asp	Asp	Ala	Asp	Phe	Val	Asp	Val	Leu	His	Thr	Tyr	Thr	Arg		
		210				215							220				
TCC	TTC	GGC	TTG	AGC	ATT	GGT	ATT	CAG	ATG	CCT	GTG	GGC	CAC	ATT	GAC	720	
Ser	Phe	Gly	Leu	Ser	Ile	Gly	Ile	Gln	Met	Pro	Val	Gly	His	Ile	Asp		
					230					235					240		
ATC	TAC	CCC	AAT	GGG	GGT	GAC	TTC	CAG	CCA	GGC	TGT	GGA	CTC	AAC	GAT	768	
Ile	Tyr	Pro	Asn	Gly	Gly	Asp	Phe	Gln	Pro	Gly	Cys	Gly	Leu	Asn	Asp		
				245					250					255			
GTC	TTG	GGA	TCA	ATT	GCA	TAT	GGA	ACA	ATC	ACA	GAG	GTG	GTA	AAA	TGT	816	
Val	Leu	Gly	Ser	Ile	Ala	Tyr	Gly	Thr	Ile	Thr	Glu	Val	Val	Lys	Cys		
			260					265						270			
GAG	CAT	GAG	CGA	GCC	GTC	CAC	CTC	TTT	GTT	GAC	TCT	CTG	GTG	AAT	CAG	864	
Glu	His	Glu	Arg	Ala	Val	His	Leu	Phe	Val	Asp	Ser	Leu	Val	Asn	Gln		
		275					280						285				
GAC	AAG	CCG	AGT	TTT	GCC	TTC	CAG	TGC	ACT	GAC	TCC	AAT	CGC	TTC	AAA	912	
Asp	Lys	Pro	Ser	Phe	Ala	Phe	Gln	Cys	Thr	Asp	Ser	Asn	Arg	Phe	Lys		
		290				295							300				
AAG	GGG	ATC	TGT	CTG	AGC	TGC	CGC	AAG	AAC	CGT	TGT	AAT	AGC	ATT	GGC	960	
Lys	Gly	Ile	Cys	Leu	Ser	Cys	Arg	Lys	Asn	Arg	Cys	Asn	Ser	Ile	Gly		
		305			310					315					320		
TAC	AAT	GCC	AAG	AAA	ATG	AGG	AAC	AAG	AGG	AAC	AGC	AAA	ATG	TAC	CTA	1008	
Tyr	Asn	Ala	Lys	Lys	Met	Arg	Asn	Lys	Arg	Asn	Ser	Lys	Met	Tyr	Leu		
				325					330					335			
AAA	ACC	CGG	GCA	GGC	ATG	CCT	TTC	AGA	GGT	AAC	CTT	CAG	TCC	CTG	GAG	1056	

7'
Cont.

Lys Thr Arg Ala Gly Met Pro Phe Arg Gly Asn Leu Gln Ser Leu Glu
 340 345 350
 TGT CCC TGA
 Cys Pro

1065

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 354 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Ser Asn Ser Val Pro Leu Leu Cys Phe Trp Ser Leu Cys Tyr Cys
 1 5 10 15
 Phe Ala Ala Gly Ser Pro Val Pro Phe Gly Pro Glu Gly Arg Leu Glu
 20 25 30
 Asp Lys Leu His Lys Pro Lys Ala Thr Gln Thr Glu Val Lys Pro Ser
 35 40 45
 Val Arg Phe Asn Leu Arg Thr Ser Lys Asp Pro Glu His Glu Gly Cys
 50 55 60
 Tyr Leu Ser Val Gly His Ser Gln Pro Leu Glu Asp Cys Ser Phe Asn
 65 70 75 80
 Met Thr Ala Lys Thr Phe Phe Ile Ile His Gly Trp Thr Met Ser Gly
 85 90 95
 Ile Phe Glu Asn Trp Leu His Lys Leu Val Ser Ala Leu His Thr Arg
 100 105 110
 Glu Lys Asp Ala Asn Val Val Val Asp Trp Leu Pro Leu Ala His
 115 120 125
 Gln Leu Tyr Thr Asp Ala Val Asn Asn Thr Arg Val Val Gly His Ser
 130 135 140
 Ile Ala Arg Met Leu Asp Trp Leu Gln Glu Lys Asp Asp Phe Ser Leu
 145 150 155 160
 Gly Asn Val His Leu Ile Gly Tyr Ser Leu Gly Ala His Val Ala Gly
 165 170 175
 Tyr Ala Gly Asn Phe Val Lys Gly Thr Val Gly Arg Ile Thr Gly Leu
 180 185 190
 Asp Pro Ala Gly Pro Met Phe Glu Gly Ala Asp Ile His Lys Arg Leu
 195 200 205
 Ser Pro Asp Asp Ala Asp Phe Val Asp Val Leu His Thr Tyr Thr Arg
 210 215 220
 Ser Phe Gly Leu Ser Ile Gly Ile Gln Met Pro Val Gly His Ile Asp
 225 230 235 240
 Ile Tyr Pro Asn Gly Asp Phe Gln Pro Gly Cys Gly Leu Asn Asp
 245 250 255
 Val Leu Gly Ser Ile Ala Tyr Gly Thr Ile Thr Glu Val Val Lys Cys
 260 265 270
 Glu His Glu Arg Ala Val His Leu Phe Val Asp Ser Leu Val Asn Gln
 275 280 285
 Asp Lys Pro Ser Phe Ala Phe Gln Cys Thr Asp Ser Asn Arg Phe Lys
 290 295 300
 Lys Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn Ser Ile Gly
 305 310 315 320
 Tyr Asn Ala Lys Lys Met Arg Asn Lys Arg Asn Ser Lys Met Tyr Leu
 325 330 335
 Lys Thr Arg Ala Gly Met Pro Phe Arg Gly Asn Leu Gln Ser Leu Glu
 340 345 350
 Cys Pro

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2565 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 252..1754

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAATTCGCGG	CCGCGTCGAC	GGCGGCTCAG	GACGAGGGCA	GATCTCGTTC	TGGGGCAAGC	60
CGTTGACACT	CGCTCCCTGC	CACCGCCCGG	GCTCCGTGCC	GCCAAGTTTT	CATTTTCCAC	120
CTTCTCTGCC	TCCAGTCCCC	CAGCCCCTGG	CCGAGAGAAG	GGTCTTACCG	GCCGGGATTG	180
CTGGAAACAC	CAAGAGGTGG	TTTTTGTTTT	TTAAAACTTC	TGTTTCTTGG	GAGGGGGTGT	240
GGCGGGGCAG	G ATG AGC AAC TCC GTT CCT CTG CTC TGT TTC TGG AGC CTC	290				
	Met Ser Asn Ser Val Pro Leu Leu Cys Phe Trp Ser Leu					
	1 5 10					
TGC TAT TGC TTT GCT GCG GGG AGC CCC GTA CCT TTT GGT CCA GAG GGA	338					
Cys Tyr Cys Phe Ala Ala Gly Ser Pro Val Pro Phe Gly Pro Glu Gly						
15 20 25						
CGG CTG GAA GAT AAG CTC CAC AAA CCC AAA GCT ACA CAG ACT GAG GTC	386					
Arg Leu Glu Asp Lys Leu His Lys Pro Lys Ala Thr Gln Thr Glu Val						
30 35 40 45						
AAA CCA TCT GTG AGG TTT AAC CTC CGC ACC TCC AAG GAC CCA GAG CAT	434					
Lys Pro Ser Val Arg Phe Asn Leu Arg Thr Ser Lys Asp Pro Glu His						
50 55 60						
GAA GGA TGC TAC CTC TCC GTC GGC CAC AGC CAG CCC TTA GAA GAC TGC	482					
Glu Gly Cys Tyr Leu Ser Val Gly His Ser Gln Pro Leu Glu Asp Cys						
65 70 75						
AGT TTC AAC ATG ACA GCT AAA ACC TTT TTC ATC ATT CAC GGA TGG ACG	530					
Ser Phe Asn Met Thr Ala Lys Thr Phe Phe Ile Ile His Gly Trp Thr						
80 85 90						
ATG AGC GGT ATC TTT GAA AAC TGG CTG CAC AAA CTC GTG TCA GCC CTG	578					
Met Ser Gly Ile Phe Glu Asn Trp Leu His Lys Leu Val Ser Ala Leu						
95 100 105						
CAC ACA AGA GAG AAA GAC GCC AAT GTA GTT GTG GTT GAC TGG CTC CCC	626					
His Thr Arg Glu Lys Asp Ala Asn Val Val Val Asp Trp Leu Pro						
110 115 120 125						
CTG GCC CAC CAG CTT TAC ACG GAT GCG GTC AAT AAT ACC AGG GTG GTG	674					
Leu Ala His Gln Leu Tyr Thr Asp Ala Val Asn Asn Thr Arg Val Val						
130 135 140						
GGA CAC AGC ATT GCC AGG ATG CTC GAC TGG CTG CAG GAG AAG GAC GAT	722					
Gly His Ser Ile Ala Arg Met Leu Asp Trp Leu Gln Glu Lys Asp Asp						
145 150 155						
TTT TCT CTC GGG AAT GTC CAC TTG ATC GGC TAC AGC CTC GGA GCG CAC	770					
Phe Ser Leu Gly Asn Val His Leu Ile Gly Tyr Ser Leu Gly Ala His						
160 165 170						
GTG GCC GGG TAT GCA GGC AAC TTC GTG AAA GGA ACG GTG GGC CGA ATC	818					
Val Ala Gly Tyr Ala Gly Asn Phe Val Lys Gly Thr Val Gly Arg Ile						
175 180 185						
ACA GGT TTG GAT CCT GCC GGG CCC ATG TTT GAA GGG GCC GAC ATC CAC	866					
Thr Gly Leu Asp Pro Ala Gly Pro Met Phe Glu Gly Ala Asp Ile His						
190 195 200 205						
AAG AGG CTC TCT CCG GAC GAT GCA GAT TTT GTG GAT GTC CTC CAC ACC	914					
Lys Arg Leu Ser Pro Asp Asp Ala Asp Phe Val Asp Val Leu His Thr						
210 215 220						
TAC ACG CGT TCC TTC GGC TTG AGC ATT GGT ATT CAG ATG CCT GTG GGC	962					
Tyr Thr Arg Ser Phe Gly Leu Ser Ile Gly Ile Gln Met Pro Val Gly						
225 230 235						
CAC ATT GAC ATC TAC CCC AAT GGG GGT GAC TTC CAG CCA GGC TGT GGA	1010					

7'
cont.

His	Ile	Asp	Ile	Tyr	Pro	Asn	Gly	Gly	Asp	Phe	Gln	Pro	Gly	Cys	Gly		
		240					245				250						
CTC	AAC	GAT	GTC	TTG	GGA	TCA	ATT	GCA	TAT	GGA	ACA	ATC	ACA	GAG	GTG	1058	
Leu	Asn	Asp	Val	Leu	Gly	Ser	Ile	Ala	Tyr	Gly	Thr	Ile	Thr	Glu	Val		
		255					260				265						
GTA	AAA	TGT	GAG	CAT	GAG	CGA	GCC	GTC	CAC	CTC	TTT	GTT	GAC	TCT	CTG	1106	
Val	Lys	Cys	Glu	His	Glu	Arg	Ala	Val	His	Leu	Phe	Val	Asp	Ser	Leu		
		270					275				280				285		
GTG	AAT	CAG	GAC	AAG	CCG	AGT	TTT	GCC	TTC	CAG	TGC	ACT	GAC	TCC	AAT	1154	
Val	Asn	Gln	Asp	Lys	Pro	Ser	Phe	Ala	Phe	Gln	Cys	Thr	Asp	Ser	Asn		
				290						295				300			
CGC	TTC	AAA	AAG	GGG	ATC	TGT	CTG	AGC	TGC	CGC	AAG	AAC	CGT	TGT	AAT	1202	
Arg	Phe	Lys	Lys	Gly	Ile	Cys	Leu	Ser	Cys	Arg	Lys	Asn	Arg	Cys	Asn		
			305							310				315			
AGC	ATT	GGC	TAC	AAT	GCC	AAG	AAA	ATG	AGG	AAC	AAG	AGG	AAC	AGC	AAA	1250	
Ser	Ile	Gly	Tyr	Asn	Ala	Lys	Lys	Met	Arg	Asn	Lys	Arg	Asn	Ser	Lys		
		320					325				330						
ATG	TAC	CTA	AAA	ACC	CGG	GCA	GGC	ATG	CCT	TTC	AGA	GTT	TAC	CAT	TAT	1298	
Met	Tyr	Leu	Lys	Thr	Arg	Ala	Gly	Met	Pro	Phe	Arg	Val	Tyr	His	Tyr		
		335					340				345						
CAG	ATG	AAA	ATC	CAT	GTC	TTC	AGT	TAC	AAG	AAC	ATG	GGA	GAA	ATT	GAG	1346	
Gln	Met	Lys	Ile	His	Val	Phe	Ser	Tyr	Lys	Asn	Met	Gly	Glu	Ile	Glu		
		350				355				360				365			
CCC	ACC	TTT	TAC	GTC	ACC	CTT	TAT	GGC	ACT	AAT	GCA	GAT	TCC	CAG	ACT	1394	
Pro	Thr	Phe	Tyr	Val	Thr	Leu	Tyr	Gly	Thr	Asn	Ala	Asp	Ser	Gln	Thr		
				370						375				380			
CTG	CCA	CTG	GAA	ATA	GTG	GAG	CGG	ATC	GAG	CAG	AAT	GCC	ACC	AAC	ACC	1442	
Leu	Pro	Leu	Glu	Ile	Val	Glu	Arg	Ile	Glu	Gln	Asn	Ala	Thr	Asn	Thr		
			385							390				395			
TTC	CTG	GTC	TAC	ACC	GAG	GAG	GAC	TTG	GGA	GAC	CTC	TTG	AAG	ATC	CAG	1490	
Phe	Leu	Val	Tyr	Thr	Glu	Glu	Asp	Leu	Gly	Asp	Leu	Leu	Lys	Ile	Gln		
		400					405				410						
CTC	ACC	TGG	GAG	GGG	GCC	TCT	CAG	TCT	TGG	TAC	AAC	CTG	TGG	AAG	GAG	1538	
Leu	Thr	Trp	Glu	Gly	Ala	Ser	Gln	Ser	Trp	Tyr	Asn	Leu	Trp	Lys	Glu		
		415				420					425						
TTT	CGC	AGC	TAC	CTG	TCT	CAA	CCC	CGC	AAC	CCC	GGA	CGG	GAG	CTG	AAT	1586	
Phe	Arg	Ser	Tyr	Leu	Ser	Gln	Pro	Arg	Asn	Pro	Gly	Arg	Glu	Leu	Asn		
		430				435				440				445			
ATC	AGG	CGC	ATC	CGG	GTG	AAG	TCT	GGG	GAA	ACC	CAG	CGG	AAA	CTG	ACA	1634	
Ile	Arg	Arg	Ile	Arg	Val	Lys	Ser	Gly	Glu	Thr	Gln	Arg	Lys	Leu	Thr		
				450						455				460			
TTT	TGT	ACA	GAA	GAC	CCT	GAG	AAC	ACC	AGC	ATA	TCC	CCA	GGC	CGG	GAG	1682	
Phe	Cys	Thr	Glu	Asp	Pro	Glu	Asn	Thr	Ser	Ile	Ser	Pro	Gly	Arg	Glu		
		465						470			475						
CTC	TGG	TTT	CGC	AAG	TGT	CGG	GAT	GGC	TGG	AGG	ATG	AAA	AAC	GAA	ACC	1730	
Leu	Trp	Phe	Arg	Lys	Cys	Arg	Asp	Gly	Trp	Arg	Met	Lys	Asn	Glu	Thr		
		480					485				490						
AGT	CCC	ACT	GTG	GAG	CTT	CCC	TGA	GGGTGCCCGG	GCAAGTCTTG	CCAGCAAGGC						1784	
Ser	Pro	Thr	Val	Glu	Leu	Pro											
		495				500											
AGCAAGACTT	CCTGCTATCC	AAGCCCATGG	AGGAAAGTTA	CTGCTGAGGA	CCCACCCAAT											1844	
GGAAGGATTC	TTCTCAGCCT	TGACCCTGGA	GCACTGGGAA	CAACTGGTCT	CCTGTGATGG											1904	
CTGGGACTCC	TCGCGGGAGG	GGACTGCGCT	GCTATAGCTC	TTGCTGCCTC	TCTTGAATAG											1964	
CTCTAACTCC	AAACCTCTGT	CCACACCTCC	AGAGCACCAA	GTCCAGATTT	GTGTGTAAGC											2024	
AGCTGGGTGC	CTGGGGCCTC	TCGTGCACAC	TGGATTGGTT	TCTCAGTTGC	TGGGCGAGCC											2084	
TGTACTCTGC	CTGACGAGGA	ACGCTGGCTC	CGAAGAGGCC	CTGTGTAGAA	GGCTGTGTCAG											2144	
TGCTCAGCCT	GCTTTGAGCC	TCAGTGAGAA	GTCCTTCCGA	CAGGAGCTGA	CTCATGTCAG											2204	
GATGGCAGGC	CTGGTATCTT	GCTCGGGCCC	TGGCTGTTGG	GGTTCTCATG	GGTTGCACTG											2264	
ACCATACTGC	TTACGTCTTA	GCCATTCCGT	CCTGCTCCCC	AGCTCACTCT	CTGAAGCACA											2324	
CATCATTTGGC	TTTCTATTTT	TTCTGTTCAT	TTTTTAATTG	AGCAAATGTC	TATTGAACAC											2384	
TTAAAATTAA	TTAGAAATGTG	GTAATGGACA	TATTACTGAG	CCTCTCCATT	TGGAACCCAG											2444	

7' cont.

TGGAGTTGGG ATTTCTAGAC CCTCTTTCTG TTTGGATGGT GTATGTGTAT ATGCATGGGG	2504
AAAGGCACCT GGGGCCTGGG GGAGGCTATA GGATATAAGC AGTCGACGCG GCCGCGAATT	2564
C	2565

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 500 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

7' cont.

Met	Ser	Asn	Ser	Val	Pro	Leu	Leu	Cys	Phe	Trp	Ser	Leu	Cys	Tyr	Cys
1				5					10					15	
Phe	Ala	Ala	Gly	Ser	Pro	Val	Pro	Phe	Gly	Pro	Glu	Gly	Arg	Leu	Glu
			20					25					30		
Asp	Lys	Leu	His	Lys	Pro	Lys	Ala	Thr	Gln	Thr	Glu	Val	Lys	Pro	Ser
		35					40				45				
Val	Arg	Phe	Asn	Leu	Arg	Thr	Ser	Lys	Asp	Pro	Glu	His	Glu	Gly	Cys
	50					55				60					
Tyr	Leu	Ser	Val	Gly	His	Ser	Gln	Pro	Leu	Glu	Asp	Cys	Ser	Phe	Asn
65					70				75					80	
Met	Thr	Ala	Lys	Thr	Phe	Phe	Ile	Ile	His	Gly	Trp	Thr	Met	Ser	Gly
			85						90					95	
Ile	Phe	Glu	Asn	Trp	Leu	His	Lys	Leu	Val	Ser	Ala	Leu	His	Thr	Arg
			100					105					110		
Glu	Lys	Asp	Ala	Asn	Val	Val	Val	Val	Asp	Trp	Leu	Pro	Leu	Ala	His
		115					120					125			
Gln	Leu	Tyr	Thr	Asp	Ala	Val	Asn	Asn	Thr	Arg	Val	Val	Gly	His	Ser
	130					135					140				
Ile	Ala	Arg	Met	Leu	Asp	Trp	Leu	Gln	Glu	Lys	Asp	Asp	Phe	Ser	Leu
145				150					155					160	
Gly	Asn	Val	His	Leu	Ile	Gly	Tyr	Ser	Leu	Gly	Ala	His	Val	Ala	Gly
			165					170					175		
Tyr	Ala	Gly	Asn	Phe	Val	Lys	Gly	Thr	Val	Gly	Arg	Ile	Thr	Gly	Leu
			180				185						190		
Asp	Pro	Ala	Gly	Pro	Met	Phe	Glu	Gly	Ala	Asp	Ile	His	Lys	Arg	Leu
	195					200					205				
Ser	Pro	Asp	Asp	Ala	Asp	Phe	Val	Asp	Val	Leu	His	Thr	Tyr	Thr	Arg
	210					215					220				
Ser	Phe	Gly	Leu	Ser	Ile	Gly	Ile	Gln	Met	Pro	Val	Gly	His	Ile	Asp
225				230					235					240	
Ile	Tyr	Pro	Asn	Gly	Gly	Asp	Phe	Gln	Pro	Gly	Cys	Gly	Leu	Asn	Asp
			245					250					255		
Val	Leu	Gly	Ser	Ile	Ala	Tyr	Gly	Thr	Ile	Thr	Glu	Val	Val	Lys	Cys
			260					265					270		
Glu	His	Glu	Arg	Ala	Val	His	Leu	Phe	Val	Asp	Ser	Leu	Val	Asn	Gln
	275					280					285				
Asp	Lys	Pro	Ser	Phe	Ala	Phe	Gln	Cys	Thr	Asp	Ser	Asn	Arg	Phe	Lys
	290					295					300				
Lys	Gly	Ile	Cys	Leu	Ser	Cys	Arg	Lys	Asn	Arg	Cys	Asn	Ser	Ile	Gly
305				310					315					320	
Tyr	Asn	Ala	Lys	Lys	Met	Arg	Asn	Lys	Arg	Asn	Ser	Lys	Met	Tyr	Leu
			325					330					335		
Lys	Thr	Arg	Ala	Gly	Met	Pro	Phe	Arg	Val	Tyr	His	Tyr	Gln	Met	Lys
			340					345					350		
Ile	His	Val	Phe	Ser	Tyr	Lys	Asn	Met	Gly	Glu	Ile	Glu	Pro	Thr	Phe
	355					360					365				
Tyr	Val	Thr	Leu	Tyr	Gly	Thr	Asn	Ala	Asp	Ser	Gln	Thr	Leu	Pro	Leu
	370					375					380				

Glu Ile Val Glu Arg Ile Glu Gln Asn Ala Thr Asn Thr Phe Leu Val

385 390 395 400
 Tyr Thr Glu Glu Asp Leu Gly Asp Leu Leu Lys Ile Gln Leu Thr Trp
 405 410 415
 Glu Gly Ala Ser Gln Ser Trp Tyr Asn Leu Trp Lys Glu Phe Arg Ser
 420 425 430
 Tyr Leu Ser Gln Pro Arg Asn Pro Gly Arg Glu Leu Asn Ile Arg Arg
 435 440 445
 Ile Arg Val Lys Ser Gly Glu Thr Gln Arg Lys Leu Thr Phe Cys Thr
 450 455 460
 Glu Asp Pro Glu Asn Thr Ser Ile Ser Pro Gly Arg Glu Leu Trp Phe
 465 470 475 480
 Arg Lys Cys Arg Asp Gly Trp Arg Met Lys Asn Glu Thr Ser Pro Thr
 485 490 495
 Val Glu Leu Pro
 500

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1035 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1035

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

7' *cont*
 ATG AGC AAC TCC GTT CCT CTG CTC TGT TTC TGG AGC CTC TGC TAT TGC 48
 Met Ser Asn Ser Val Pro Leu Leu Cys Phe Trp Ser Leu Cys Tyr Cys
 1 5 10 15
 TTT GCT GCG GGG AGC CCC GTA CCT TTT GGT CCA GAG GGA CGG CTG GAA 96
 Phe Ala Ala Gly Ser Pro Val Pro Phe Gly Pro Glu Gly Arg Leu Glu
 20 25 30
 GAT AAG CTC CAC AAA CCC AAA GCT ACA CAG ACT GAG GTC AAA CCA TCT 144
 Asp Lys Leu His Lys Pro Lys Ala Thr Gln Thr Glu Val Lys Pro Ser
 35 40 45
 GTG AGG TTT AAC CTC CGC ACC TCC AAG GAC CCA GAG CAT GAA GGA TGC 192
 Val Arg Phe Asn Leu Arg Thr Ser Lys Asp Pro Glu His Glu Gly Cys
 50 55 60
 TAC CTC TCC GTC GGC CAC AGC CAG CCC TTA GAA GAC TGC AGT TTC AAC 240
 Tyr Leu Ser Val Gly His Ser Gln Pro Leu Glu Asp Cys Ser Phe Asn
 65 70 75 80
 ATG ACA GCT AAA ACC TTT TTC ATC ATT CAC GGA TGG ACG ATG AGC GGT 288
 Met Thr Ala Lys Thr Phe Phe Ile Ile His Gly Trp Thr Met Ser Gly
 85 90 95
 ATC TTT GAA AAC TGG CTG CAC AAA CTC GTG TCA GCC CTG CAC ACA AGA 336
 Ile Phe Glu Asn Trp Leu His Lys Leu Val Ser Ala Leu His Thr Arg
 100 105 110
 GAG AAA GAC GCC AAT GTA GTT GTG GTT GAC TGG CTC CCC CTG GCC CAC 384
 Glu Lys Asp Ala Asn Val Val Val Val Asp Trp Leu Pro Leu Ala His
 115 120 125
 CAG CTT TAC ACG GAT GCG GTC AAT AAT ACC AGG GTG GTG GGA CAC AGC 432
 Gln Leu Tyr Thr Asp Ala Val Asn Asn Thr Arg Val Val Gly His Ser
 130 135 140
 ATT GCC AGG ATG CTC GAC TGG CTG CAG GAG AAG GAC GAT TTT TCT CTC 480
 Ile Ala Arg Met Leu Asp Trp Leu Gln Glu Lys Asp Asp Phe Ser Leu
 145 150 155 160
 GGG AAT GTC CAC TTG ATC GGC TAC AGC CTC GGA GCG CAC GTG GCC GGG 528
 Gly Asn Val His Leu Ile Gly Tyr Ser Leu Gly Ala His Val Ala Gly
 165 170 175
 TAT GCA GGC AAC TTC GTG AAA GGA ACG GTG GGC CGA ATC ACA GGT TTG 576

7' cont.

Tyr	Ala	Gly	Asn	Phe	Val	Lys	Gly	Thr	Val	Gly	Arg	Ile	Thr	Gly	Leu	
			180						185					190		
GAT	CCT	GCC	GGG	CCC	ATG	TTT	GAA	GGG	GCC	GAC	ATC	CAC	AAG	AGG	CTC	624
Asp	Pro	Ala	Gly	Pro	Met	Phe	Glu	Gly	Ala	Asp	Ile	His	Lys	Arg	Leu	
		195					200					205				
TCT	CCG	GAC	GAT	GCA	GAT	TTT	GTG	GAT	GTC	CTC	CAC	ACC	TAC	ACG	CGT	672
Ser	Pro	Asp	Asp	Ala	Asp	Phe	Val	Asp	Val	Leu	His	Thr	Tyr	Thr	Arg	
	210					215					220					
TCC	TTC	GGC	TTG	AGC	ATT	GGT	ATT	CAG	ATG	CCT	GTG	GGC	CAC	ATT	GAC	720
Ser	Phe	Gly	Leu	Ser	Ile	Gly	Ile	Gln	Met	Pro	Val	Gly	His	Ile	Asp	
	225				230					235					240	
ATC	TAC	CCC	AAT	GGG	GGT	GAC	TTC	CAG	CCA	GGC	TGT	GGA	CTC	AAC	GAT	768
Ile	Tyr	Pro	Asn	Gly	Gly	Asp	Phe	Gln	Pro	Gly	Cys	Gly	Leu	Asn	Asp	
			245					250					255			
GTC	TTG	GGA	TCA	ATT	GCA	TAT	GGA	ACA	ATC	ACA	GAG	GTG	GTA	AAA	TGT	816
Val	Leu	Gly	Ser	Ile	Ala	Tyr	Gly	Thr	Ile	Thr	Glu	Val	Val	Lys	Cys	
			260					265					270			
GAG	CAT	GAG	CGA	GCC	GTC	CAC	CTC	TTT	GTT	GAC	TCT	CTG	GTG	AAT	CAG	864
Glu	His	Glu	Arg	Ala	Val	His	Leu	Phe	Val	Asp	Ser	Leu	Val	Asn	Gln	
		275					280					285				
GAC	AAG	CCG	AGT	TTT	GCC	TTC	CAG	TGC	ACT	GAC	TCC	AAT	CGC	TTC	AAA	912
Asp	Lys	Pro	Ser	Phe	Ala	Phe	Gln	Cys	Thr	Asp	Ser	Asn	Arg	Phe	Lys	
	290					295					300					
AAG	GGG	ATC	TGT	CTG	AGC	TGC	CGC	AAG	AAC	CGT	TGT	AAT	AGC	ATT	GGC	960
Lys	Gly	Ile	Cys	Leu	Ser	Cys	Arg	Lys	Asn	Arg	Cys	Asn	Ser	Ile	Gly	
	305				310					315				320		
TAC	AAT	GCC	AAG	AAA	ATG	AGG	AAC	AAG	AGG	AAC	AGC	AAA	ATG	TAC	CTA	1008
Tyr	Asn	Ala	Lys	Lys	Met	Arg	Asn	Lys	Arg	Asn	Ser	Lys	Met	Tyr	Leu	
			325					330					335			
AAA	ACC	CGG	GCA	GGC	ATG	CCT	TTC	AGA								1035
Lys	Thr	Arg	Ala	Gly	Met	Pro	Phe	Arg								
			340					345								

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 345 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met	Ser	Asn	Ser	Val	Pro	Leu	Leu	Cys	Phe	Trp	Ser	Leu	Cys	Tyr	Cys	
1				5					10					15		
Phe	Ala	Ala	Gly	Ser	Pro	Val	Pro	Phe	Gly	Pro	Glu	Gly	Arg	Leu	Glu	
		20						25					30			
Asp	Lys	Leu	His	Lys	Pro	Lys	Ala	Thr	Gln	Thr	Glu	Val	Lys	Pro	Ser	
	35					40						45				
Val	Arg	Phe	Asn	Leu	Arg	Thr	Ser	Lys	Asp	Pro	Glu	His	Glu	Gly	Cys	
	50					55					60					
Tyr	Leu	Ser	Val	Gly	His	Ser	Gln	Pro	Leu	Glu	Asp	Cys	Ser	Phe	Asn	
	65			70					75					80		
Met	Thr	Ala	Lys	Thr	Phe	Phe	Ile	Ile	His	Gly	Trp	Thr	Met	Ser	Gly	
			85						90				95			
Ile	Phe	Glu	Asn	Trp	Leu	His	Lys	Leu	Val	Ser	Ala	Leu	His	Thr	Arg	
		100						105					110			
Glu	Lys	Asp	Ala	Asn	Val	Val	Val	Val	Asp	Trp	Leu	Pro	Leu	Ala	His	
	115					120						125				
Gln	Leu	Tyr	Thr	Asp	Ala	Val	Asn	Asn	Thr	Arg	Val	Val	Gly	His	Ser	
	130					135					140					
Ile	Ala	Arg	Met	Leu	Asp	Trp	Leu	Gln	Glu	Lys	Asp	Asp	Phe	Ser	Leu	
	145			150						155				160		
Gly	Asn	Val	His	Leu	Ile	Gly	Tyr	Ser	Leu	Gly	Ala	His	Val	Ala	Gly	

Tyr Ala Gly Asn Phe Val Lys Gly Thr Val Gly Arg Ile Thr Gly Leu
 165 170 175
 180 185 190
 Asp Pro Ala Gly Pro Met Phe Glu Gly Ala Asp Ile His Lys Arg Leu
 195 200 205
 Ser Pro Asp Asp Ala Asp Phe Val Asp Val Leu His Thr Tyr Thr Arg
 210 215 220
 Ser Phe Gly Leu Ser Ile Gly Ile Gln Met Pro Val Gly His Ile Asp
 225 230 235 240
 Ile Tyr Pro Asn Gly Asp Phe Gln Pro Gly Cys Gly Leu Asn Asp
 245 250 255
 Val Leu Gly Ser Ile Ala Tyr Gly Thr Ile Thr Glu Val Val Lys Cys
 260 265 270
 Glu His Glu Arg Ala Val His Leu Phe Val Asp Ser Leu Val Asn Gln
 275 280 285
 Asp Lys Pro Ser Phe Ala Phe Gln Cys Thr Asp Ser Asn Arg Phe Lys
 290 295 300
 Lys Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn Ser Ile Gly
 305 310 315 320
 Tyr Asn Ala Lys Lys Met Arg Asn Lys Arg Asn Ser Lys Met Tyr Leu
 325 330 335
 Lys Thr Arg Ala Gly Met Pro Phe Arg
 340 345

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 225 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..225

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CTG GGA TCC ATC GCC TAT GGC ACG ATC GCG GAG GTG GTG AAG TGC GAG 48
 Leu Gly Ser Ile Ala Tyr Gly Thr Ile Ala Glu Val Val Lys Cys Glu
 1 5 10 15
 CAT GAG CGG GCC GTG CAT CTC TTT GTG GAC TCC CTG GTG AAC CAG GAC 96
 His Glu Arg Ala Val His Leu Phe Val Asp Ser Leu Val Asn Gln Asp
 20 25 30
 AAG CCG AGC TTT GCC TTC CAG TGC ACA GAC TCC AAC CGC TTC AAA AAA 144
 Lys Pro Ser Phe Ala Phe Gln Cys Thr Asp Ser Asn Arg Phe Lys Lys
 35 40 45
 GGG ATC TGT CTC AGC TGC CGG AAG AAC CGC TGT AAC GGC ATC GGC TAC 192
 Gly Ile Cys Leu Ser Cys Arg Lys Asn Arg Cys Asn Gly Ile Gly Tyr
 50 55 60
 AAT GCT AAG AAG ACG AGG AAT AAG AGG AAC ACC 225
 Asn Ala Lys Lys Thr Arg Asn Lys Arg Asn Thr
 65 70 75

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 75 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Leu Gly Ser Ile Ala Tyr Gly Thr Ile Ala Glu Val Val Lys Cys Glu
 1 5 10 15
 His Glu Arg Ala Val His Leu Phe Val Asp Ser Leu Val Asn Gln Asp

f'
cont.

(i) SEQUENCE CHARACTERISTICS:

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

Met 1	Glu	Ser	Lys	Ala 5	Leu	Leu	Val	Leu	Thr 10	Leu	Ala	Val	Trp	Leu 15	Gln
Ser	Leu	Thr	Ala 20	Ser	Arg	Gly	Gly	Val 25	Ala	Ala	Ala	Asp	Gln 30	Arg	Arg
Asp	Phe	Ile	Asp 35	Ile	Glu	Ser	Lys 40	Phe	Ala	Leu	Arg	Thr 45	Pro	Glu	Asp
Thr	Ala 50	Glu	Asp	Thr	Cys	His 55	Leu	Ile	Pro	Gly	Val 60	Ala	Glu	Ser	Val
Ala 65	Thr	Cys	His	Phe	Asn 70	His	Ser	Ser	Lys	Thr 75	Phe	Met	Val	Ile	His 80
Gly	Trp	Thr	Val 85	Thr	Gly	Met	Tyr	Glu	Ser 90	Trp	Val	Pro	Lys	Leu 95	Val
Ala	Ala	Leu	Tyr 100	Lys	Arg	Glu	Pro	Asp 105	Ser	Asn	Val	Ile	Val 110	Val	Asp
Trp	Leu	Ser	Arg 115	Ala	Gln	Glu	His 120	Tyr	Pro	Val	Ser	Ala 125	Gly	Tyr	Thr
Lys	Leu 130	Val	Gly	Gln	Asp	Val 135	Ala	Arg	Phe	Ile	Asn 140	Trp	Met	Glu	Glu
Glu 145	Phe	Asn	Tyr	Pro	Leu 150	Asp	Asn	Val	His	Leu 155	Leu	Gly	Tyr	Ser	Leu 160
Gly	Ala	His	Ala 165	Ala	Gly	Ile	Ala	Gly	Ser 170	Leu	Thr	Asn	Lys	Lys 175	Val
Asn	Arg	Ile	Thr 180	Gly	Leu	Asp	Pro	Ala 185	Gly	Pro	Asn	Phe	Glu 190	Tyr	Ala
Glu	Ala	Pro	Ser 195	Arg	Leu	Ser	Pro 200	Asp	Asp	Ala	Asp	Phe 205	Val	Asp	Val
Leu	His 210	Thr	Phe	Thr	Arg	Gly 215	Ser	Pro	Gly	Arg	Ser 220	Ile	Gly	Ile	Gln
Lys 225	Pro	Val	Gly	His	Val 230	Asp	Ile	Tyr	Pro	Asn 235	Gly	Gly	Thr	Phe	Gln 240
Pro	Gly	Cys	Asn 245	Ile	Gly	Glu	Ala	Ile	Arg 250	Val	Ile	Ala	Glu	Arg 255	Gly
Leu	Gly	Asp	Val 260	Asp	Gln	Leu	Val	Lys 265	Cys	Ser	His	Glu	Arg 270	Ser	Ile
His	Leu	Phe 275	Ile	Asp	Ser	Leu	Leu 280	Asn	Glu	Glu	Asn	Pro 285	Ser	Lys	Ala
Tyr	Arg 290	Cys	Ser	Ser	Lys	Glu 295	Ala	Phe	Glu	Lys	Gly 300	Leu	Cys	Leu	Ser
Cys 305	Arg	Lys	Asn	Arg	Cys 310	Asn	Asn	Leu	Gly	Tyr 315	Glu	Ile	Asn	Lys	Val 320
Arg	Ala	Lys	Arg 325	Ser	Ser	Lys	Met	Tyr 330	Leu	Lys	Thr	Arg	Ser	Gln 335	Met
Pro	Tyr	Lys	Val 340	Phe	His	Tyr	Gln	Val 345	Lys	Ile	His	Phe	Ser 350	Gly	Thr
Glu	Ser	Glu	Thr	His	Thr	Asn	Gln	Ala	Phe	Glu	Ile	Ser	Leu	Tyr	Gly

	355		360		365										
Thr	Val	Ala	Glu	Ser	Glu	Asn	Ile	Pro	Phe	Thr	Leu	Pro	Glu	Val	Ser
	370					375					380				
Thr	Asn	Lys	Thr	Tyr	Ser	Phe	Leu	Ile	Tyr	Thr	Glu	Val	Asp	Ile	Gly
385					390					395				400	
Glu	Leu	Leu	Met	Leu	Lys	Leu	Lys	Trp	Lys	Ser	Asp	Ser	Tyr	Phe	Ser
			405						410					415	
Trp	Ser	Asp	Trp	Trp	Ser	Ser	Pro	Gly	Phe	Ala	Ile	Gln	Lys	Ile	Arg
			420					425					430		
Val	Lys	Ala	Gly	Glu	Thr	Gln	Lys	Lys	Val	Ile	Phe	Cys	Ser	Arg	Glu
		435					440					445			
Lys	Val	Ser	His	Leu	Gln	Lys	Gly	Lys	Ala	Pro	Ala	Val	Phe	Val	Lys
	450					455					460				
Cys	His	Asp	Lys	Ser	Leu	Asn	Lys	Lys	Ser	Gly					
465					470					475					

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 499 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met	Asp	Thr	Ser	Pro	Leu	Cys	Phe	Ser	Ile	Leu	Leu	Val	Leu	Cys	Ile
1				5				10					15		
Phe	Ile	Gln	Ser	Ala	Leu	Gly	Gln	Ser	Leu	Lys	Pro	Glu	Pro	Phe	
			20				25					30			
Gly	Arg	Arg	Ala	Gln	Ala	Val	Glu	Thr	Asn	Lys	Thr	Leu	His	Glu	Met
			35				40					45			
Lys	Thr	Arg	Phe	Leu	Leu	Phe	Gly	Glu	Thr	Asn	Gln	Gly	Cys	Gln	Ile
	50					55					60				
Arg	Ile	Asn	His	Pro	Asp	Thr	Leu	Gln	Glu	Cys	Gly	Phe	Asn	Ser	Ser
65					70					75				80	
Leu	Pro	Leu	Val	Met	Ile	Ile	His	Gly	Trp	Ser	Val	Asp	Gly	Val	Leu
			85					90					95		
Glu	Asn	Trp	Ile	Trp	Gln	Met	Val	Ala	Leu	Lys	Ser	Gln	Pro	Ala	
			100					105				110			
Gln	Pro	Val	Asn	Val	Gly	Leu	Val	Asp	Trp	Ile	Thr	Leu	Ala	His	Asp
		115					120					125			
His	Tyr	Thr	Ile	Ala	Val	Arg	Asn	Thr	Arg	Leu	Val	Gly	Lys	Glu	Val
	130					135					140				
Ala	Ala	Leu	Leu	Arg	Trp	Leu	Glu	Glu	Ser	Val	Gln	Leu	Ser	Arg	Ser
145					150					155				160	
His	Val	His	Leu	Ile	Gly	Tyr	Ser	Leu	Gly	Ala	His	Val	Ser	Gly	Phe
			165						170					175	
Ala	Gly	Ser	Ser	Ile	Gly	Gly	Thr	His	Lys	Ile	Gly	Arg	Ile	Thr	Gly
			180					185					190		
Leu	Asp	Ala	Ala	Gly	Pro	Leu	Phe	Glu	Gly	Ser	Ala	Pro	Ser	Asn	Arg
		195					200					205			
Leu	Ser	Pro	Asp	Asp	Ala	Asn	Phe	Val	Asp	Ala	Ile	His	Thr	Phe	Thr
	210					215					220				
Arg	Glu	His	Met	Gly	Leu	Ser	Val	Gly	Ile	Lys	Gln	Pro	Ile	Gly	His
225					230					235				240	
Tyr	Asp	Phe	Tyr	Pro	Asn	Gly	Gly	Ser	Phe	Gln	Pro	Gly	Cys	His	Phe
			245						250				255		
Leu	Glu	Leu	Tyr	Arg	His	Ile	Ala	Gln	His	Gly	Phe	Asn	Ala	Ile	Thr
			260					265					270		
Gln	Thr	Ile	Lys	Cys	Ser	His	Glu	Arg	Ser	Val	His	Leu	Phe	Ile	Asp
	275						280					285			
Ser	Leu	Leu	His	Ala	Gly	Thr	Gln	Ser	Met	Ala	Tyr	Pro	Cys	Gly	Asp

71
cont.

(2) INFORMATION FOR SEQ ID NO:15:

(A) LENGTH: 465 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Asp Pro Ala Glu Pro Cys Phe Gln Gly Thr Pro Glu Leu Val Arg Leu

[illegible]

(2) INFORMATION FOR SEQ ID NO:18:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 10 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:
GATCAATCGC

10

(2) INFORMATION FOR SEQ ID NO:19:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
TAGGACATGC ACAGTGTAACT CTG

23

71
cont.
(2) INFORMATION FOR SEQ ID NO:20:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
GATTGTGCTG GCCACTTCTC

20

(2) INFORMATION FOR SEQ ID NO:21:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 19 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:
GACTCTCCAG GGACTGAAG

19

(2) INFORMATION FOR SEQ ID NO:22:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 48 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
 (ix) FEATURE:
 (A) NAME/KEY: modified_base
 (B) LOCATION: 36
 (D) OTHER INFORMATION: /mod_base= i
 (ix) FEATURE:
 (A) NAME/KEY: modified_base
 (B) LOCATION: 37
 (D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
 (A) NAME/KEY: modified_base
 (B) LOCATION: 41
 (D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
 (A) NAME/KEY: modified_base
 (B) LOCATION: 42
 (D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
 (A) NAME/KEY: modified_base
 (B) LOCATION: 46
 (D) OTHER INFORMATION: /mod_base= i

(ix) FEATURE:
 (A) NAME/KEY: modified_base
 (B) LOCATION: 47
 (D) OTHER INFORMATION: /mod_base= i

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:
 CUACUACUAC UAGGCCACGC GTCGACTAGT ACGGGNNGG NNGGGNNG 48

(2) INFORMATION FOR SEQ ID NO:23:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 28 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:
 CACACACAGG CCACGCGTCG ACTAGTAC 28

7'
 6' 5' 3' 2' 1'
 (2) INFORMATION FOR SEQ ID NO:24:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:
 ACCACCATGG AGAGCAAAGC CCTG 24

(2) INFORMATION FOR SEQ ID NO:25:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:
 CCAGTTTCAG CCTGACTTCT TATTC 25

(2) INFORMATION FOR SEQ ID NO:26:
 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:
GGCTGTGGAC TCAACGATGT C 21

(2) INFORMATION FOR SEQ ID NO:27:
(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 22 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:
CCGGGTGGGT AGGTACATTT TG 22

(2) INFORMATION FOR SEQ ID NO:28:
(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:
GGGGGTGACT TCCAGCCAGG CTGTG 25

7' cord.
(2) INFORMATION FOR SEQ ID NO:29:
(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
AACTCTGAAA GGCATGCCTG CCCGG 25

(2) INFORMATION FOR SEQ ID NO:30:
(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:
TGAAGGTCGG AGTCAACGGA TTTGGT 26

(2) INFORMATION FOR SEQ ID NO:31:
(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "Oligonucleotide"
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:
CATGTGGGCC ATGAGGTCCA CCAC 24